BNL Role in ATLAS Computing

DOE Annual HEP Program Review

S. Rajagopalan

April 22, 2003

BROOKHAVEN
NATIONAL LABORATORY

ATLAS Computing

❖ **Software**

- Core Software
- Reconstruction & Simulation

❖ **Analysis**

- Physics capabilities at the LHC
- Preparation for Analysis

❖ **Facility support**

- Local (U.S.) software support
- BNL Tier 1 Center
- Production (Simulation & Reconstruction)
- Develop & deploy Grid Tools & Services

Background

❖ **Pre-1999 era:**

- Geant3 based simulation
- Fortran based reconstruction software
- Physics Technical Design Report published in 1999
  - Detector & Physics Performance

❖ **Post-1999 era:**

- Work on Geant 4 based simulation
- Object Oriented Reconstruction software
  - Core Services, sub-system software & combined reconstruction
  - Adopted the LHCb framework as a starting point
  - The LAr Reconstruction was the first client of the new framework.
    - Considerable influence in core design aspects
- Data Challenges + Physics studies with the new software
  - Data Challenge 1 in progress

ATLAS Core Software Breakdown

❖ Framework

✔ Event Data Model ⎤
                      ⎬ Areas where BNL plays
✔ Data Management   ⎦   a significant role

❖ Detector Description

❖ Graphics

❖ Analysis Tools

❖ Grid Integration

EDM Coordinator: S. Rajagopalan
Raw Data Model Coordinator: H. Ma
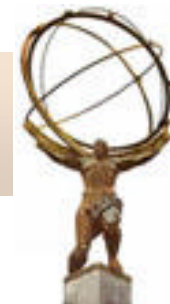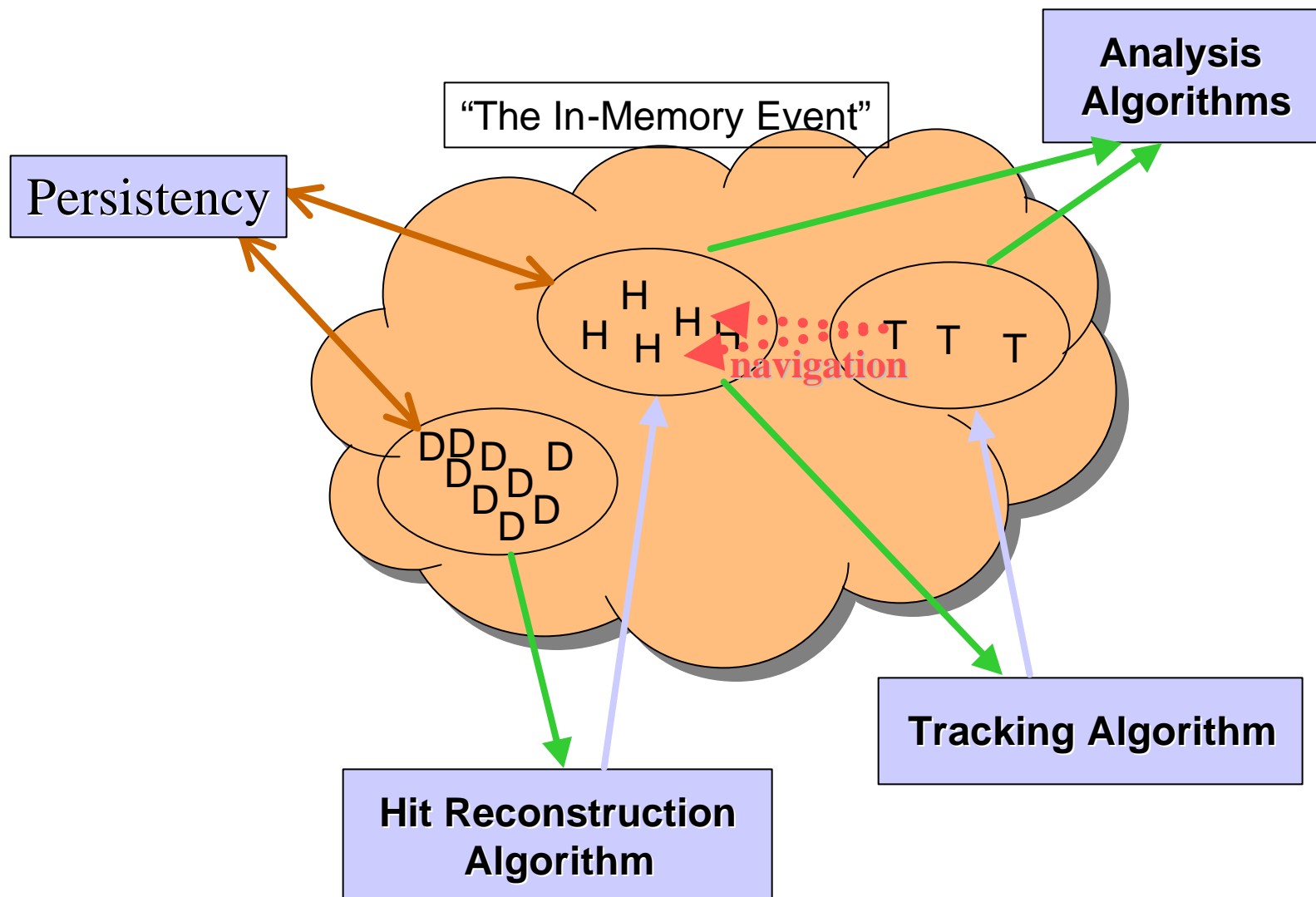LCG Applications Area Coordinator: T. Wenaus

Architecture-Team (A-Team):   S. Rajagopalan, T. Wenaus

Event Data Model

❖ **Major contributions to EDM infrastructure from:**

  ▪ H. Ma, S. Rajagopalan

❖ **EDM Infrastructure (StoreGate) provides tools for:**

  ▪ Memory management for Event Data Objects

  ▪ On-demand access to objects from persistency

  ▪ Persistable navigation between objects (Track $\rightarrow$ Hits)

  ▪ History (how objects are created)

  ▪ Efficient access to data in regions of interest

❖ **ATLAS has baselined StoreGate as its choice of the EDM infrastructure.**

  ▪ High Level Trigger (HLT) will also use StoreGate.

BROOKHAVEN
NATIONAL LABORATORY

Event Data Model Infrastructure (StoreGate)

Analysis
Algorithms

"The In-Memory Event"

Persistency

H
H H H H H
navigation
T T T

D D D D
D D D
D D D

Tracking Algorithm

Hit Reconstruction
Algorithm

BROOKHAVEN
NATIONAL LABORATORY

Event Data Model (2)

❖ (H. Ma, S. Rajagopalan)

❖ BNL contribution to EDM software for physicists:

- Design of the overall ATLAS Raw Data Flow
    - Implementation of the Calorimeter aspects
    - Emulation of ByteStream data flow and efficient unpacking and access in regions of interest
        - Used in High Level Trigger
    - Coordination of the Raw Data Flow across sub-sytems

- Overall coordination of the ATLAS EDM effort
    - Design & implementation of Calorimeter EDM
    - Several aspects of the EDM for combined reconstruction

Persistency Support

❖ The baseline persistent technology for ATLAS changed from Objectivity to ROOT in 2001.

- All LHC experiments have a common baseline.

- BNL has significant expertise in this area (leveraging efforts from RHIC)

- We immediately provided an interim ROOT based persistency mechanism for ATLAS that is being widely used in several applications.

- Long term efforts are closely integrated with the LCG (LHC Computing & Grid project) to provide a common persistency solution.

BROOKHAVEN
NATIONAL LABORATORY

The LCG Project

❖ To help the LHC experiments prepare, build and operate the computing environment needed to manage and analyze multi-PB scale data coming from each detector.

❖ Main working body of the LCG is the Project Execution Board (PEB) chaired by Les Robertson.

❖ PEB has four areas, each with a project manager:

✓ Applications   (T. Wenaus)

  Grid Technology

  Fabrics

  Grid Deployment

BROOKHAVEN
NATIONAL LABORATORY

LCG Applications Area

❖ Identifies and provides solutions for common projects:

  ▪ Long term advantages in providing resources, support and maintenance

❖ Identified Common Projects are:

  ✔ Persistency Framework Project (POOL)

  ▪ Physicist Interface Project (PI)
    ❑ Interfaces and tools with which physicists will use the software

  ▪ Core Libraries and Services Project (SEAL)
    ❑ Core libraries, Object Dictionary, Scripting Services etc.

  ▪ Software Process & Infrastructure Project (SPI)
    ❑ Provides basic environment and tools for software development

  ▪ Simulation Project
    ❑ Support for Geant3, Geant4, Detector Description

BROOKHAVEN
NATIONAL LABORATORY

Persistency Support (2)

❖ **BNL current efforts are to integrate the ATLAS software with POOL to provide persistency support.**
  - Its longer term include direct contributions to the POOL effort in the areas of Persistency Support and Event Collections
  - D. Adams, V. Fine, H. Ma, V. Perevotchikov

❖ **BNL is also providing MySQL based persistency for conditions data.** (S. Kanadasamy, H. Ma, A. Undrus)
  - Currently in use by ATLAS (especially by Calorimeter)
  - While we may switch to a new conditions database, the technology independent interface to the algorithms will be stable.

❖ **Hong Ma: Overall Liquid Argon database coordinator**
  - Oversees the needs of production, installation, online, offline and testbeam efforts in the LAr community.
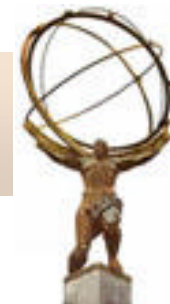
Magda

❖ **Distributed Data Manager for cataloging and data replication**

  ▪ Developed by Wengsheng Deng (BNL).

❖ **Heavily used in ATLAS Data Challenges (DC0 & DC1)**

  ▪ Catalog of ATLAS data at Alberta, CERN, Lyon, INFN (CNAF, Milan), FZK, IFIC, IHEP.su.itep.ru, NorduGrid, RAL and many US institutions.

  ▪ 288K files in catalog with total size of 77.5 TB as of 2003-04-15

  ▪ Main component in US testbed production

  ▪ Data Replication task has transferred 10 TB between BNL HPSS and CERN Castor

❖ **Tested in EDG testbed and demonstrated to be useful.**

❖ **It will be implemented as a file catalog back end to the LCG POOL persistency framework.**

BROOKHAVEN
NATIONAL LABORATORY

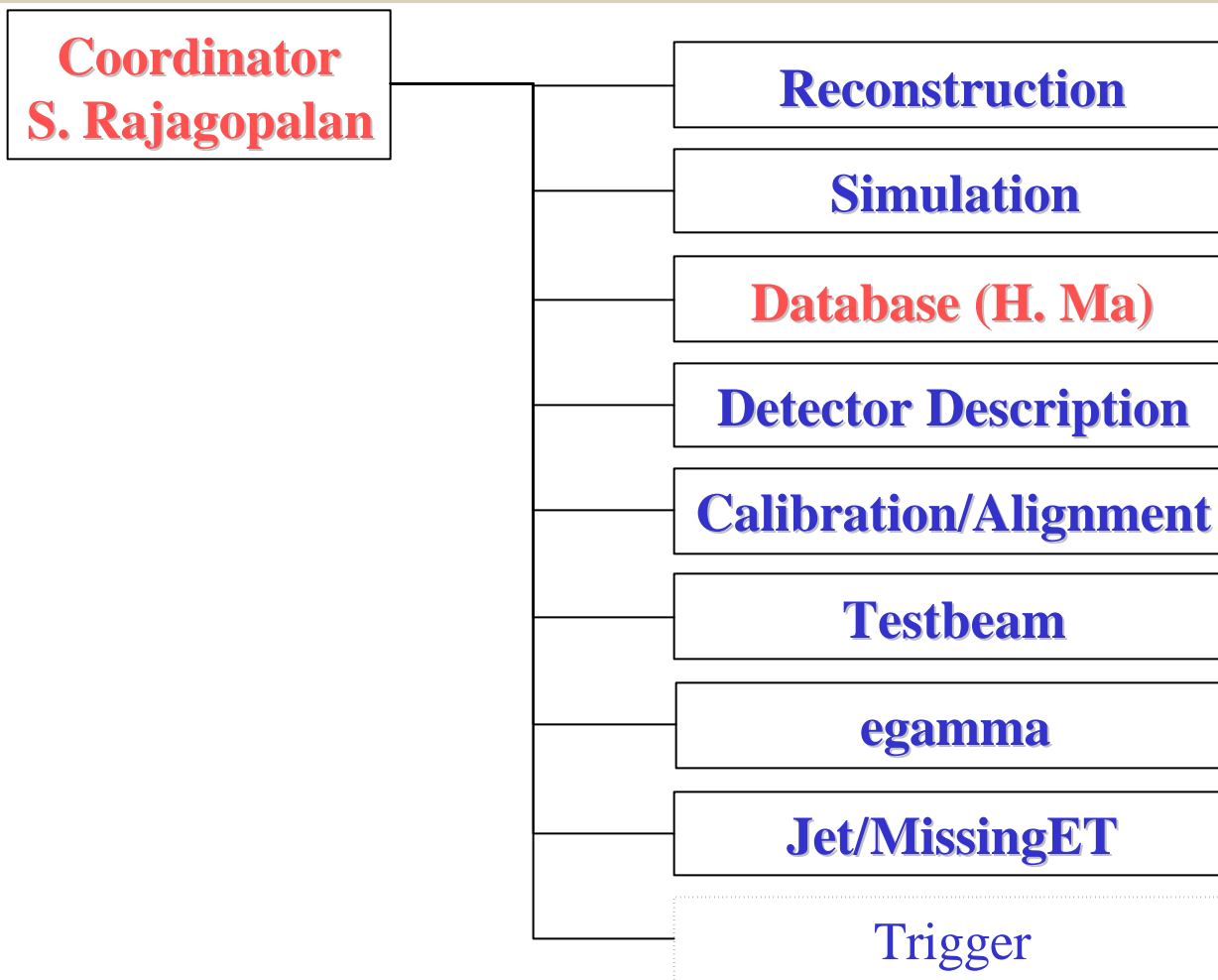# Simulation, Reconstruction & Analysis

Geant3 Simulation

❖ P. Nevski serves as the ATLAS Geant3 simulation coordinator.

- Data Challenge 0 & 1 utilize Geant3 based simulation.
- Recent work geared toward implementing updated geometry.

- Geant4 based simulation expected to be used for DC2 (April 2004)
  - Expect continued maintenance and support of Geant3 for ~ 2 years.
    - Until G4 has been validated.

BROOKHAVEN
NATIONAL LABORATORY

Liquid Argon Software Organization

Coordinator S. Rajagopalan	Reconstruction
	Simulation
	Database (H. Ma)
	Detector Description
	Calibration/Alignment
	Testbeam
	egamma
	Jet/MissingET
	Trigger

➤ Current efforts focussed on delivering the needed software for Data Challenge 2 & Combined Testbeam run: Spring 2004

Liquid Argon Reconstruction

❖ **Significant contributions in the Liquid Argon sub-system:**

- ■ (H. Ma, F. Paige, S. Rajagopalan, K. Yip)

- ■ Responsible for deployment of the full sequence of steps, developing algorithms and associated EDM:

    - ❑ Digitization & ROD Emulation

    - ❑ Cell, Tower and Cluster Reconstruction

        - – Implementation and coordination of calibration/corrections

    - ❑ Combined Reconstruction to identify eγ, τ and Jet candidates

❖ **Used our success in influencing the overall architectural design and the design of reconstruction software in other sub-systems.**

Other Reconstruction Activities

❖ **Muon Reconstruction** (K. Assamagan, Y. Fisyak, D. Adams)

- EDM and Data Converters for Muon Reconstruction

❖ **Combined Reconstruction Activities:** (H. Ma, F. Paige, S. Rajagopalan)

- e-gamma Reconstruction
    - Develop algorithms to identify egamma candidates
- Tau reconstruction
- Jet Reconstruction and Hadronic Calibration
- Missing E_T Reconstruction

❖ **Physics capabilities in the SUSY sector** (F. Paige)

- F. Paige serves as the co-coordinator of the ATLAS SUSY Group
- GMSB & mSUGRA models have been explored
- Easy to discover if it exists, challenge is to understand underlying model
- Analysis being redone with full simulation + new software

BROOKHAVEN
NATIONAL LABORATORY

Jet Calibration

❖ **ATLAS calorimeters non-compensating.**

  ▪ Hadronic showers less dense than EM : so weight cells more

  ▪ H1 calibration (at cell level) produces improved linearity, jet and MissingET resolution compared to standard sampling calibration

$$71.6\% / \sqrt{E_T} \oplus 3.9\%$$

BROOKHAVEN
NATIONAL LABORATORY
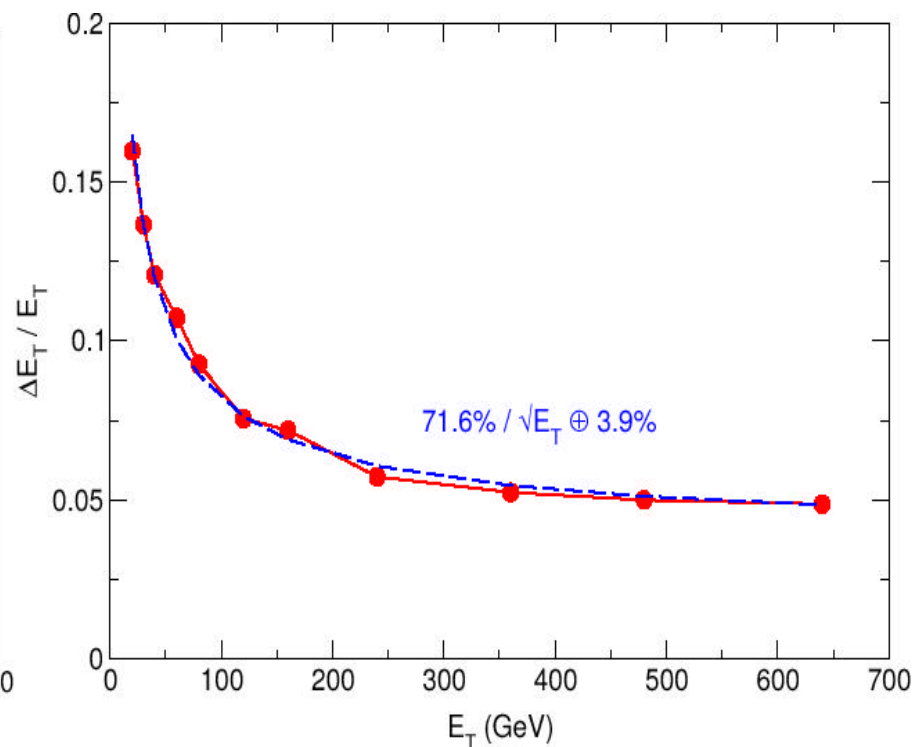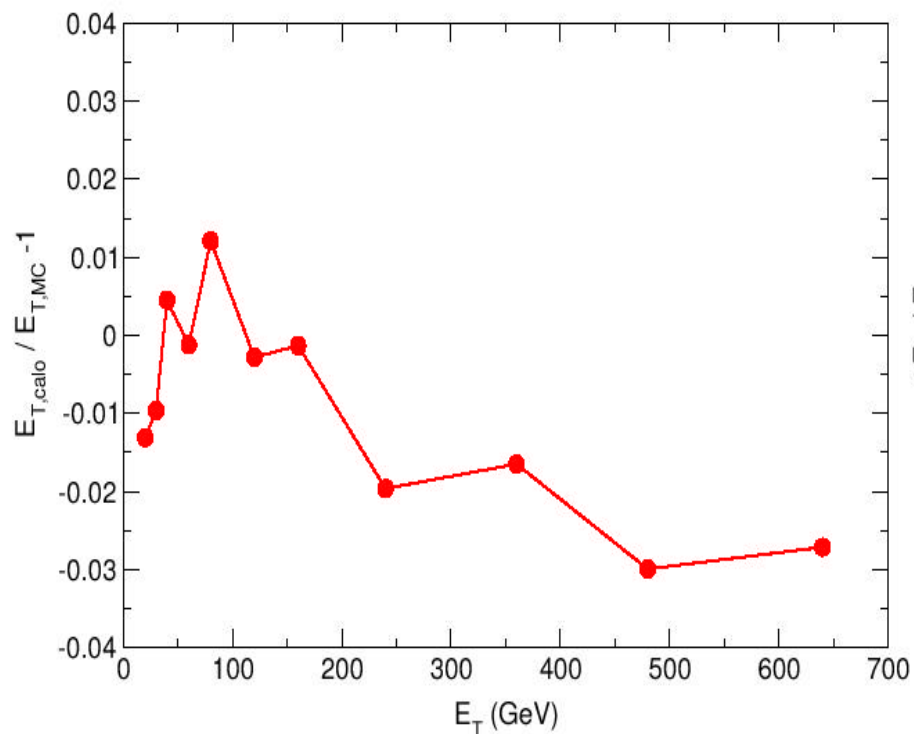
Tau Reconstruction

❖ **Important ingredient in SUSY models.**

  ▪ **Must rely on hadronic decays; hence large QCD backgrouds.**

Select narrow jets using many shape variables in likelihood function.

A cut on this likelihood function + imposing track requirements: Have achieved S/B ~ 3 at 35% efficiency for P_T > 35 GeV
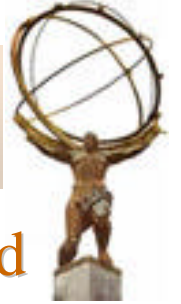
Data corresponds to ~ 2 fb^{-1} for a specific mSugra point
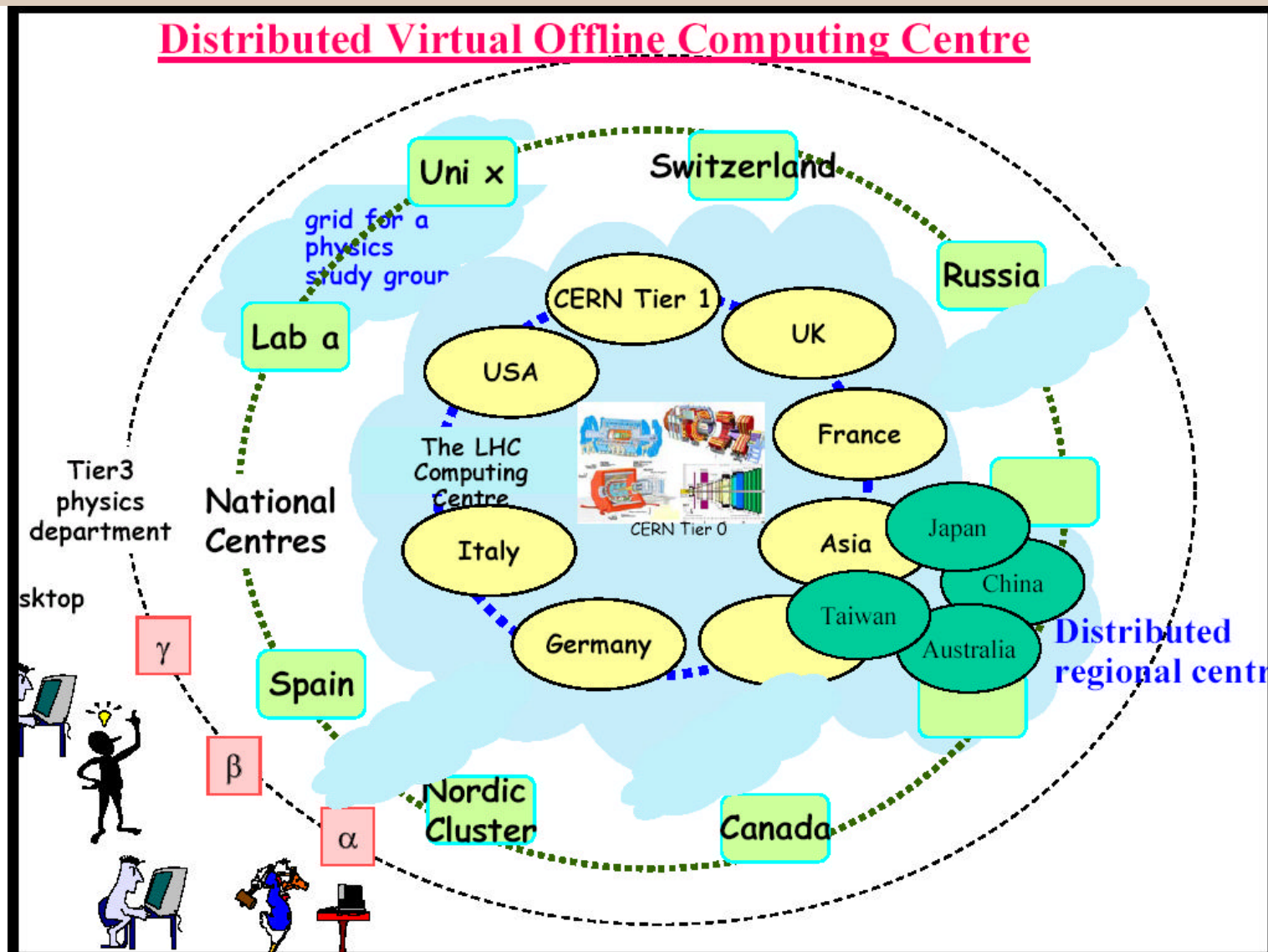
BROOKHAVEN
NATIONAL LABORATORY
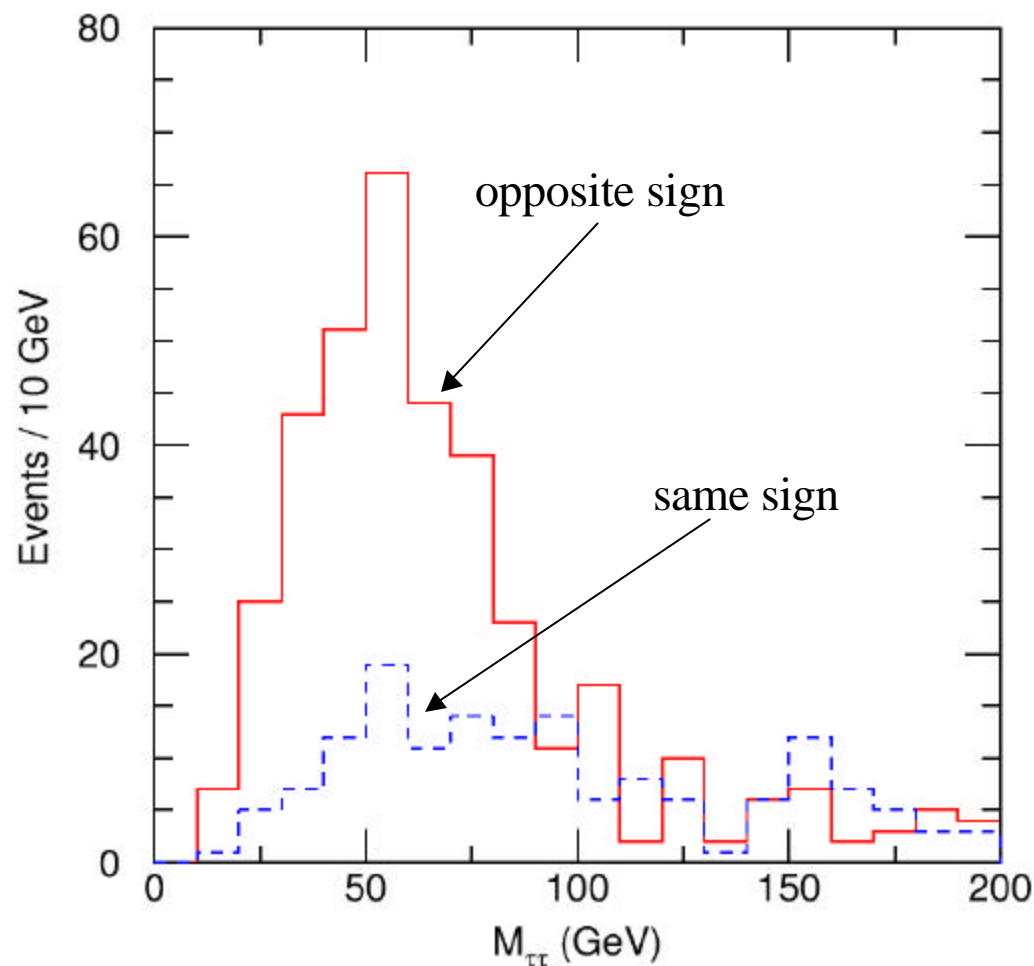
Software Support, Facilities, Production & Grid

Software Support

❖ A full time librarian (A. Undrus) to manage the ATLAS and associated software locally at BNL for U.S.

  ▪ Evaluation of release tools and their subsequent deployment

❖ Alex is a member of the ATLAS software infrastructure team (SIT) representing U.S. ATLAS.

❖ BNL developed the nightly-build system that is now deployed at CERN.

  ▪ Central tool in day to day work for software developers
  ▪ Testing procedures developed and integrated in automated builds
  ▪ System likely to be deployed in the LCG Application area as well.

BROOKHAVEN
NATIONAL LABORATORY

Distributed Virtual Offline Computing Centre

BROOKHAVEN
NATIONAL LABORATORY

Tier 1 Center

❖ **BNL has been selected as the U.S. Tier 1 Facility**

- B. Gibbard (manager), R. Baker (deputy) + staff
- Currently operational at ~ 1% of the required 2008 capacity
- Total capacity ~20% of the Tier0 center at CERN
- Five Tier 2 centers elsewhere in U.S., each with 20% of Tier1 capacity

	Tape Based Model	3 Center Disk Model	Standalone Disk Model
CPU (kSPECint95)	209	329	500
Disk (TBytes)	365	483	1000
Tape (PBytes)	2	2	2
Disk (GBytes/sec)	10	20	20
Tape (MBytes/sec)	1000	200	200
WAN (Mbit/sec)	4610	9115	9115
		1/3+1/6 of ESD on disk	Add other 2/3 of ESD
ESD pass each month		ESD pass each day	

BROOKHAVEN
NATIONAL LABORATORY

US ATLAS Regional Center (Tier 1) at BNL

November 2002

11.4 TBytes FibreChannel RAID Disk

NFS

NFS

SAN Switch

E480 (NFS Server)

E450

RCF Backup Server

WAN at OC12

62 Intel/Linux
Dual 700/450 MHz
256/512 MBytes
9/18 GBytes
100 Mbit Ethernet
(3,200 SPECint95)
Processing Farm

Dual Intel

Dual Intel

LAN Switch

Gigabit Connections to Data Servers &WAN

RCF Switch/ Firewall

RCF HPSS CORE Server

RCF Tape Silo

US ATLAS HPSS Server

9940B Tape Drive

6 Intel/Linux
Dual 1 GHz
1 GBytes
for
Sftwr Devel &
US ATLAS Lib

Dual Intel

Dual Intel

AFS Servers

Cache

Intel
Infrastructure Servers

Web & Email Servers

NIS/DNS/DTS Servers

SSH & Grid Gateways

~150 GBytes RAID Disk

AFS

US ATLAS Equipment

RCF Infrastructure

BROOKHAVEN
NATIONAL LABORATORY

ATLAS Data Challenges

❖ A series of Data Challenges with increasing complexity is planned in ATLAS.

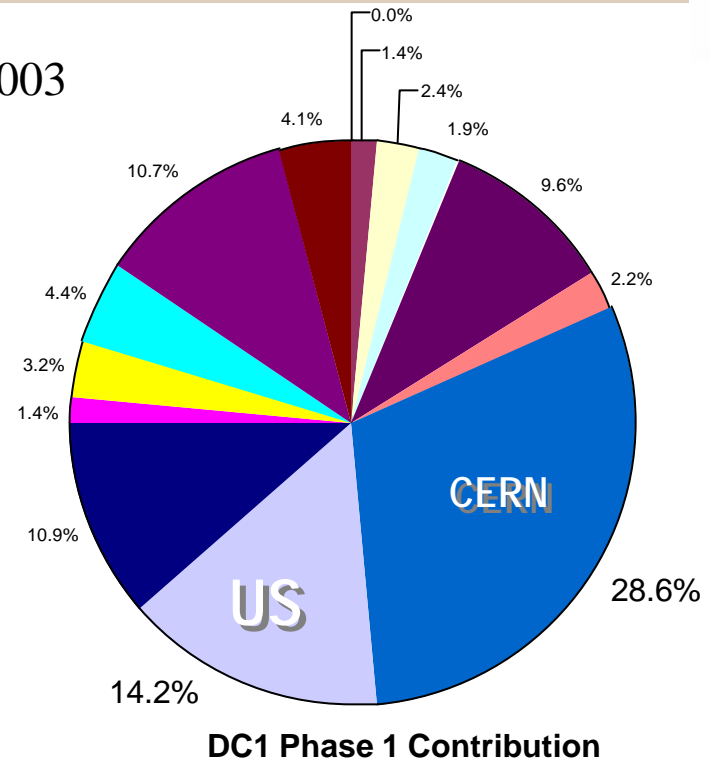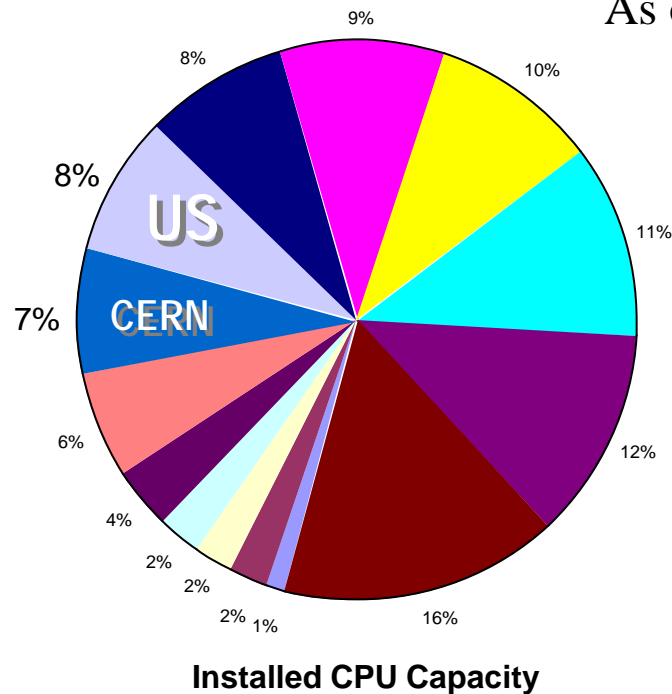❖ Goal is to help us prepare by exercising the computing infrastructure, the complete software and study the physics

DC0:	**~10^5 events,**	**December 2002**
DC1:	**~10^7 events,**	**January – June 2003**
DC2:	**~10^8 events,**	**Spring 2004**
DC3:	**~5 x 10^8 events,**	**2005**
DC4:	**~10^9 events,**	**2006**

❖ BNL has played a vital role in the DC0 & DC1 production

- ▪ P. Nevski (simulation), Y. Fisyak (Reconstruction) + …

BROOKHAVEN
NATIONAL LABORATORY

As of Feb. 2003

Installed CPU Capacity

DC1 Phase 1 Contribution

❖ BNL serves as one of the handful of primary data repositories.

- Consequently a primary analysis site

❖ DC1 Phase 2 underway, which includes:

- Pile-up production, Reconstruction production

Grid Tools & Services

❖ This effort is largely a PPDG funded project:

- MAGDA: (W. Deng)
    - Distributed manager for cataloging and data replication (W. Deng)

- DIAL: (D. Adams)
    - Distributed Interactive Analysis of Large datasets (D. Adams)

- GUMS: (T. Wlodek, D. Yu)
    - Grid User Management System
        - Virtual Organization user management
        - Site authentication & authorization software

- Monitoring: (J. Smith, D. Yu)
    - Local Cluster Monitoring integrated into Grid Middleware

- Grid oversight and liaison activities (R. Baker)

Conclusion

❖ **BNL is significantly contributing in several areas of ATLAS computing:**
  - Major responsibilities both in ATLAS and U.S. ATLAS Computing
  - As a Tier 1 Center, BNL is the main node in U.S. Grid testbed

❖ **BNL is developing significant expertise in many areas of software:**
  - In turn, this expertise allows us to rapidly start looking at physics
  - And play an important coordination role in U.S. ATLAS

❖ **Our main problem: We are understaffed.**
  - Lost Kin Yip, who was making significant contributions
  - Significant cuts in project funding as well

BROOKHAVEN
NATIONAL LABORATORY